DOCUMENT RESUME

ED 429 125 TM 029 668

AUTHOR	Barnette, J. Jackson
TITLE	Likert Response Alternative Direction: SA to SD or SD to SA: Does It Make a Difference?
PUB DATE	1999-04-00
NOTE	12p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23, 1999).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*College Students; *High School Students; High Schools; Higher Education; *Likert Scales; *Responses; Surveys; Teachers; *Test Construction; Test Items; Test Reliability
IDENTIFIERS	*Item Wording

ABSTRACT

A 20-item survey was designed in 4 forms with response set direction as "strongly disagree" (SD) to "strongly agree" (SA) and SA to SD crossed with the absence or presence of negatively worded item stems. The primary research question related to finding a primacy effect when comparing the two response direction formats. Surveys were administered, randomly by form, to 586 high school students, undergraduate and graduate students, and teachers. There were no differences in internal consistency reliability, total score means, total score variance, item means, item standard deviations, or item-to-total correlations. The presence or absence of negatively worded stems did not affect any of the survey statistics except Cronbach's alpha where there was about a 0.1 lower value when mixed stems were used (half positive and half negative) compared with having all positive or direct stems. However, there was no interaction effect with the response direction variable. It is concluded that for this survey there was no primacy effect. Reasons are proposed for why such an effect was not observed here when others have detected such an effect. These relate to differences in types of surveys, the focus of the survey, and the relationship of the topic and the respondent and how they may influence the likelihood of observing a primacy effect. (Contains 4 tables and 13 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 429 125

Likert Response Alternative Direction: SA to SD or SD to SA: Does it Make a Difference?

A paper presented at:

The 1999 Annual Meeting

of the

American Educational Research Association

April 23,1999 – Montreal

By

J. Jackson Barnette
Associate Professor
Preventive Medicine and Environmental Health
The University of Iowa
1-204 Medical Education Building
Iowa City. IA 52242
(319) 335 8905
jack-barnette@uiowa.edu

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. J. Barnette

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM029668

Running head: Response order

BEST COPY AVAILABLE

ABSTRACT

A 20-item survey was designed in four forms with response set direction as SD to SA and SA to SD crossed with the absence or presence of negatively worded item stems. The primary research question related to finding a primacy effect when comparing the two response direction formats. Surveys were administered, randomly by form, to 586 subjects. There were no differences in internal consistency reliability, total score means, total score variance, item means, item standard deviations, or item-to-total correlations. The presence or absence of negatively worded stems did not affect any of the survey statistics except Cronbach's alpha, where there was about a .1 lower value when mixed stems were used (half positive and half negative) compared with having all positive or direct stems. However, there was no interaction effect with the response direction variable. It is concluded that for this survey there was no primacy effect. Reasons why such an effect was not observed here when others have detected such an effect are proposed. These relate to differences in types of surveys, the focus of the survey, and the relationship of the topic and the respondent and how they may influence the likelihood of observing a primacy effect.

While it has not been one of the burning issues in survey design, answering the question of: "Should I use a Likert response alternative pattern that goes SA A N D SD or one that goes SD D N A SA?" reflects a practical issue that probably comes up every day. The traditional direction has been SA to SD. A search of the literature and review of the most popular texts on educational measurement provides little guidance for making such a decision. The available research has examined what is referred to as a primacy effect or the tendency for subjects to select options closer to the left side of the responses. In probably the earliest example of research on this topic, Matthews (1929) concluded that respondents were more likely to select response options to the left rather than the right on a printed survey. Carp (1974) found respondents tended to select responses presented first in an interview situation. The research of others (Johnson, 1981; Powers, Morrow, Goudy, and Keith, 1977) has not generally supported the presence of a primacy effect. Only two recent empirical studies were found (Chan, 1991 and Albanese, Prucha, Barnet, and Gjerde, 1997) where self-administered ordered-response surveys were used.

Chan (1991) administered five items from the Personal Distress (PD) Scale, a subscale of the Interpersonal Reactivity Index (Davis, 1980) to the same subjects five weeks apart with the first administration using a positive-first response alternative and the second administration using a negative-first response alternative. The alternatives used were variations on "describes me" rather than SD to SA options. Chan found there was a tendency for respondents to have higher scores when the positive-first response set was used and there were also differences in factor structures between the data sets generated with the two forms of the instrument.

Albanese, Prucha, Barnet, and Gjerde (1997) used six variations of a student evaluation of instruction form in a medical education setting. The six forms came from crossing the number of response alternatives of five, six, or seven with the response alternative pattern have the "strongly agree" option first or last. They found forms with the most positive statement first (to the left) had more positive ratings and less variance. Of course these statistics are not totally independent when a closed scale is used. As an item mean gets closer to a limit, the variance is constrained.

Neither of these studies, looking at primacy effects, examined possible interaction effects of using all direct or mixed stems or personal characteristics of respondents. Vacha-Haase (1998) has pointed out the need for examining score reliability across different studies.

The research reported here provides additional empirical evidence for answering the question and examination of possible related variables. More specifically, the research questions are:

> Are there any differences in internal consistency reliability, total score mean and variance, item means and variances, and item-total correlations between the two response directions?

Are any such differences related to the presence or absence of negatively-worded stems or respondent characteristics of sex, handedness orientation, and/or age, singularly or in interaction?

Clearly the researcher would have no control of personal characteristics of respondents and it's not likely that different versions of surveys would be developed such as a form labeled for use only by "left-handed females above 40 years of age." However, for the purpose of satisfying the inquisitive nature of this researcher, they have been included in this research. Had any of these variables been significantly related, there would be great concern and interest in finding out why that may have happened. The use of mixed item stems is certainly controlled by the researcher.

Methods

A twenty-item survey on attitude toward year-round schooling was developed and field-tested with 33 subjects. The Cronbach alpha for the original form, which used a Likert response alternative pattern of SD D N A SA, was .85. Another form was developed using the same 20 stems but a response pattern of SA A N D SD. Ten of the 20 items were randomly selected for negative wording. These ten items were negatively-worded for both of the two direction forms resulting in four forms of the instrument, a two-by-two factor pattern. One factor was response set direction (SA to SD or SD to SA). The second factor, which was crossed with the response direction factor, was the absence or presence of negatively-worded stems. Additional classification variables, which might be associated with the direction and stem type variables, were respondent sex, handedness orientation, and age. Questions were included at the end of the survey to obtain data on these variables.

Data were collected from 586 respondents who were asked to complete one of the four forms, assigned randomly. Respondents were high school students, undergraduate students, graduate students, and inservice teachers in five geographic locations in two states. Responses were converted to digits 1 to 5 and all scored in the same direction such that higher scores represented more positive agreement with the direct form of the stem. Thus, the negative stem item scores were reflected and the SA to SD responses were reflected to be in the same order as the SD to SA responses. Data were analyzed using programs from SAS® including PROC CORR, PROC UNIVARIATE, and PROC ANOVA. Inferential tests comparing Cronbach alpha values were computed using equations provided by Feldt, Woodruff, and Salih (1987) using programs developed by the author.

Results

The following table provides the results relative to the three primary dependent variables: Cronbach alpha, total mean score (sum of items) response (M_x), and standard deviation of total scores (SD_x). Although standard deviations are reported in the table, actual inferential tests of variability (Bartlett test) used variances rather than standard deviations.

There were no significant differences in alpha values when comparing them between the two response alternative directions. While it is clear the alpha values are about .1 lower when mixed stems are used compared with having all direct worded stems, this difference is not related to any type of interaction effect with the response alternative direction variable. This result is consistent with other research comparing the Cronbach alpha values in the absence and presence of negatively-worded stems (Barnette, 1997; Barnette 1999; Schriesheim & Hill, 1981).

There were no significant differences among the total score means or variances. Interaction significance tests incorporating sex, handedness orientation, and age category as factorial classification variables were all nonsignificant.

Characteristics of items were examined also including item means, standard deviations, item-to-total score correlations. Table 2 presents item means by the various configurations. Of primary interest is the comparison of the item means for the SD to SA configurations with the SA to SD configurations. There was very high consistency of the item means even when items had negative stems compared with results on those items when they had positive stems. For the entire set of 20 items the correlation between item means was .989, indicating virtually the same pattern of item means across the two response directions.

Table 3 presents item standard deviations by the various configurations. Of primary interest is the comparison of the item means for the SD to SA configurations with the SA to SD configurations. There was very high consistency of the item standard deviations even when items had negative stems compared with results on those items when they had positive stems. For the entire set of 20 items the correlation between item standard deviations was .895, indicating virtually the same pattern of item standard deviations across the two response directions.

Table 4 presents item-to-total score correlations by the various configurations. Of primary interest is the comparison of the item-to-total score correlations for the SD to SA configurations with the SA to SD configurations. There was very high consistency of the item-to-total score correlations even when items had negative stems compared with results on those items when they had positive stems. For the entire set of 20 items the correlation between item-to-total correlations was .967, indicating virtually the same pattern of item-to-total correlations across the two response directions.

Conclusions and Discussion

Based on these results, there is no evidence that the directionality of Likert response alternatives should be a concern in the design of at least some types of surveys. While this may or may not be an issue for many survey designers, it is a question frequently asked by those learning to design such surveys and perhaps is in the back of the minds of many seasoned survey designers. A primacy effect was not observed in this experiment. This indicates that at

least sometimes it may not make any difference which direction is used as related to the technical adequacy and stability of the results obtained.

In answer to the original question: "Does it make a difference?", it didn't in this situation. Not only did it not make a difference overall, the variables of whether negatively-worded items were present made no difference. In addition, personal characteristics of sex, handedness orientation, and age were not related to differences in response patterns under either response direction condition.

Why were these results not supportive of the findings of Chan and Albanese, Prucha, Barnet, and Gjerde? There are several possibilities that should be considered. There seem to be two, somewhat related, issues here that may make a difference. First, is the nature of the survey itself. The survey used in this research was attitude toward year-round schooling. The items were not ones that elicited strong responses one way or another for most respondents. The survey used in the Albanese, et al., research was a course evaluation survey, which are notorious for extreme, usually positive, responses. Surveys which have items that are likely to have responses close to one of the extremes may be more likely to be prone to a primacy effect. There may be a mental regression effect happening in these cases. There may be a difference between surveys used to collect general opinions as opposed to being used to evaluate something or someone. Related to this may be the nature of what is being evaluated. It may make a difference if the evaluation is of an inanimate object such as a product or project as opposed to a person such as a classroom instructor or supervisor.

Second, is the direct personal involvement of the respondent. In Chan's research, respondents were rating themselves on issues that could have been emotionally reactive. The nature of such self-assessment may be more prone to a primacy effect compared with assessment or attitude toward something or someone other than self. Acquiescence or the provision of socially desirable responses may also be an issue in this and similar situations.

Thus, primacy effect may be situation and assessment strategy determined rather than being determined by the nature of the instrument. In a way similar to the assessment of reliability, a survey or test in and of itself is not reliable. Only the scores generated in a specific use of the instrument possess the property of reliability (Thompson, 1994). Thus may be the case relative to primacy as well. These and other issues remain to be addressed in future research. This study should be replicated using different surveys and respondent types to confirm or refute these findings.

References

Albanese, M., Prucha, C., Barnet, J. & Gjerde, C. (1997). The direction of positive-rating placement on Likert-type items used to obtain medical students' ratings of instruction. Academic Medicine, 72, 627-630.

Barnette, J. (1997, April). Effects of item and stem reversals of survey statistics. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Barnette, J. (1999). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there's a better alternative to using those negatively-worded stems. Manuscript submitted for publication.

Carp, F. (1974). Position effects on interview responses. Journal of Gerontology, 29, 581-587.

Chan, J. (1991). Response-order effects in Likert-type scales. Educational and Psychological Measurement, 51(3), 531-540.

Davis, M. (1980). A multidimensional approach to individual differences in empathy. Catalog of selected documents in psychology, 10, 85.

Feldt, L., Woodruff, D., & Salih, F. (1987). Statistical inference for coefficient alpha. Applied Psychological Measurement, 11(1), 93-103.

Johnson, J. (1981). Effects of the order of presentation of evaluative dimensions for bipolar scales in four societies. The Journal of Social Psychology, 113, 21-27.

Matthews, C. (1929). The effect of the printed response words on an interest questionnaire. Journal of Educational Psychology, 30, 128-134.

Powers, E., Morrow, P., Goudy, W., & Keith, P. (1977). Serial order preference in survey research. Public Opinion Quarterly, 41, 80-85.

Schriesheim, C., & Hill. K. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. Educational and Psychological Measurement, 41(4), 1101-1114.

Thompson, B. (1994). Guidelines for authors. Educational and Psychological Measurement, 54, 837-847.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. Educational and Psychological Measurement, 58(1), 6-20.

Table 1

Cronbach Alpha, Means, and Standard Deviations by Survey Form

		Response alt. SD to SA	Response alt. SA to SD	Total
All positive or direct worded stems	n	146	147	293
	α	.8189	.8106	.8154
	M_X	60.185	60.939	60.563
	SD_X	9.504	9.506	9.500
Mixed stems, Half positive, Half negative	n	146	147	293
	α	.7268	.7063	.7161
	M_X	62.397	61.776	62.085
	SD_X	8.344	8.267	8.297
Total	n	292	294	586
	α	.7771	.7604	.7690
	M_X	61.291	61.357	61.324
	SD_X	8.996	8.903	8.942

Table 2. Item Means by Survey Configurations

Item	SD to SA		SA to SD		Total SD – SA n= 292	Total SA – SD n= 294	Total Positive n= 293	Total Mixed n= 293	Total n= 586
	Positive n= 146	Mixed n= 146	Positive n= 147	Mixed n= 147					
1*	3.000	2.966	3.054	2.952	2.983	3.003	3.027	2.959	2.993
2	2.432	2.514	2.279	2.476	2.473	2.378	2.355	2.495	2.425
3	3.144	3.267	3.197	3.143	3.205	3.170	3.171	3.205	3.188
4	3.116	3.144	3.116	3.204	3.130	3.160	3.116	3.174	3.145
5*	3.555	3.281	3.612	3.401	3.418	3.507	3.584	3.341	3.462
6*	4.212	4.082	4.299	4.109	4.147	4.204	4.256	4.096	4.176
7*	2.692	3.123	2.673	3.156	2.908	2.915	2.683	3.140	2.911
8*	3.096	3.158	3.122	2.939	3.127	3.031	3.109	3.048	3.078
9	3.747	3.719	3.585	3.633	3.733	3.609	3.666	3.676	3.671
10	3.500	3.630	3.482	3.483	3.565	3.483	3.491	3.556	3.524
11*	2.219	3.164	2.265	3.245	2.692	2.755	2.242	3.205	2.724
12*	2.897	3.075	2.993	3.027	2.986	3.010	2.945	3.051	2.999
13	2.514	2.603	2.558	2.490	2.558	2.524	2.536	2.546	2.541
14*	3.356	3.760	3.694	3.714	3.558	3.704	3.526	3.737	3.631
15	2.192	2.288	2.272	2.272	2.240	2.272	2.232	2.280	2.256
16	2.315	2.370	2.218	2.197	2.342	2.207	2.266	2.283	2.275
17*	2.815	2.822	2.871	2.952	2.818	2.912	2.843	2.887	2.865
18*	3.452	3.390	3.551	3.456	3.421	3.503	3.502	3.423	3.462
19	3.370	3.473	3.537	3.340	3.421	3.439	3.454	3.406	3.430
20	2.562	2.568	2.558	2.585	2.565	2.571	2.560	2.577	2.568
M_{Pos}	2.889	2.958	2.880	2.882	2.923	2.881	2.885	2.920	2.902
M_{Neg}	3.129	3.282	3.214	3.295	3.206	3.254	3.172	3.289	3.230
M_{Tot}	3.009	3.120	3.047	3.089	3.065	3.068	3.028	3.104	3.066

*indicates item had reverse wording on surveys labeled "Mixed"

Table 3 Item Standard Deviations by Survey Configurations

Item	SD to SA		SA to SD		Total SD – SA	Total SA – SD	Total Positive	Total Mixed	Total
	Positive n= 146	Mixed n= 146	Positive n= 147	Mixed n= 147	n= 292	n= 294	n= 293	n= 293	n= 586
1*	1.157	1.217	1.163	1.279	1.185	1.221	1.158	1.246	1.203
2	1.023	1.078	0.999	1.068	1.050	1.037	1.012	1.071	1.044
3	1.145	1.097	1.174	1.165	1.121	1.168	1.158	1.131	1.144
4	1.067	1.133	1.168	1.146	1.098	1.156	1.117	1.138	1.127
5*	1.051	1.113	1.050	1.121	1.089	1.089	1.049	1.116	1.089
6*	0.911	1.054	0.932	0.945	0.986	0.942	0.921	0.999	0.963
7*	0.827	0.961	0.804	0.984	0.920	0.929	0.814	0.971	0.924
8*	1.110	1.155	1.164	1.178	1.131	1.172	1.136	1.169	1.152
9	0.967	0.967	1.026	0.973	0.965	0.998	0.999	0.969	0.983
10	1.052	1.044	1.131	1.100	1.048	1.114	1.090	1.073	1.081
11*	0.928	1.297	1.009	1.236	1.222	1.229	0.968	1.266	1.225
12*	0.820	0.925	0.940	0.986	0.877	0.962	0.882	0.955	0.920
13	0.949	1.020	0.885	1.003	0.984	0.944	0.916	1.011	0.964
14*	1.068	1.059	0.955	1.014	1.081	0.983	1.025	1.035	1.035
15	0.905	1.017	0.983	1.024	0.962	1.002	0.944	1.019	0.982
16	1.106	1.057	1.004	1.070	1.081	1.036	1.055	1.065	1.060
17*	1.076	0.884	1.035	1.023	0.983	1.028	1.055	0.957	1.006
18*	0.887	0.978	0.915	0.953	0.933	0.934	0.901	0.964	0.933
19	0.947	0.919	0.788	0.969	0.933	0.887	0.873	0.945	0.909
20	1.070	1.101	1.159	1.059	1.084	1.108	1.114	1.078	1.095
M_{Pos}	1.023	1.043	1.032	1.058	1.033	1.045	1.028	1.050	1.039
M_{Neg}	0.984	1.064	0.997	1.072	1.041	1.049	0.991	1.068	1.045
M_{Tot}	1.003	1.054	1.014	1.065	1.037	1.047	1.009	1.059	1.042

*indicates item had reverse wording on surveys labeled "Mixed"

Table 4 Item Correlation with Total Score by Survey Configurations

Item	SD to SA		SA to SD		Total	Total	Total	Total	Total
	Positive n= 146	Mixed n= 146	Positive n= 147	Mixed n= 147	SD – SA n= 292	SA – SD n= 294	Positive. n= 293	Mixed n= 293	n= 586
1*	.671	.530	.561	.447	.594	.499	.616	.487	.546
2	.646	.561	.601	.611	.602	.604	.619	.586	.602
3	.718	.716	.763	.711	.717	.735	.741	.713	.726
4	.564	.511	.608	.593	.533	.601	.586	.550	.568
5*	.003	-.057	.136	-.057	-.041	.038	.071	-.059	-.001
6*	.249	.109	.017	.099	.167	.050	.133	.104	.110
7*	.604	.415	.636	.518	.515	.559	.619	.466	.537
8*	.019	.064	.029	-.082	.043	-.027	.024	-.006	.008
9	.217	.060	.233	.189	.140	.214	.221	.126	.177
10	.596	.708	.739	.669	.649	.705	.669	.689	.677
11*	.502	.184	.504	.138	.339	.303	.503	.160	.321
12*	.588	.578	.684	.489	.583	.589	.639	.532	.585
13	.490	.433	.462	.353	.461	.402	.476	.394	.432
14*	.165	-.080	-.090	-.089	.072	-.088	.050	-.084	-.004
15	.508	.408	.535	.422	.457	.478	.522	.415	.468
16	.652	.595	.637	.607	.624	.618	.641	.602	.620
17*	.591	.588	.611	.518	.586	.568	.601	.545	.576
18*	.483	.697	.530	.560	.576	.538	.508	.628	.557
19	.651	.433	.498	.527	.552	.497	.580	.482	.525
20	.564	.576	.620	.567	.564	.596	.592	.571	.580
M_{Neg}	.561	.500	.570	.525	.530	.545	.565	.513	.538
M_{Pos}	.388	.303	.362	.254	.343	.303	.376	.277	.323
M_{Tot}	.474	.401	.466	.389	.437	.424	.471	.395	.430

*indicates item had reverse wording on surveys labeled "Mixed"